SEQUENCE LISTING

(1) GENERAL INFORMATION:

  (i) APPLICANT:
    (A) NAME: Zeneca Limited
    (B) STREET: 15 Stanhope Gate
    (C) CITY: London
    (D) STATE: England
    (E) COUNTRY: United Kingdom
    (F) POSTAL CODE (ZIP): W1Y 6LN
    (G) TELEPHONE: 0171 304 5000
    (H) TELEFAX: 0171 304 5151
    (I) TELEX: 0171 304 2042

  (ii) TITLE OF INVENTION: NOVEL COMPOUNDS

  (iii) NUMBER OF SEQUENCES: 5

  (iv) COMPUTER READABLE FORM:
    (A) MEDIUM TYPE: Floppy disk
    (B) COMPUTER: IBM PC compatible
    (C) OPERATING SYSTEM: PC-DOS/MS-DOS
    (D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

  (vi) PRIOR APPLICATION DATA:
    (A) APPLICATION NUMBER: GB 9716162.4
    (B) FILING DATE: 01-AUG-1997

(2) INFORMATION FOR SEQ ID NO: 1:

  (i) SEQUENCE CHARACTERISTICS:
    (A) LENGTH: 5154 base pairs
    (B) TYPE: nucleic acid
    (C) STRANDEDNESS: single
    (D) TOPOLOGY: linear

  (ii) MOLECULE TYPE: other nucleic acid

  (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

CAAGCGCGCA ATTAACCCTC ACTAAAGGGA ACACCAACAC GTCGCCAGGA
CTGCGCCGTT   60

CGCTGCGCTC ATAGGCGGCG ATTTCATCAA GGGTGGCAAG GATCGCCTGG
TCGACGGTCA   120

GGTCGTCCTC GACGCGGTTG CCCTCCTCGT CCTGTTCCAG GGTGAGTGGG
CGATACCAGG   180

TGTCCACCGG GAAGGTACGG CCCGACACCT CGACAATCGG CGCATCGTCG
AAGTGCTTGG   240

AAAAGCGCTC CAGGTCGATG GTGGCCGAGG TGATGATGAC TTTCAGGTCG
GGGCGACGCG   300

GCAACAGGGT CTTGAGGTAG CCGAGCAGGA AGTCGATGTT CAGGCTGCGT
TCGTGGGCTT   360

CGTCGACGAC AGGCTCGCGT TATGGCTCCG CTTTCTGCGG CTCTCCTACC
CTGGCATGGT   420

GTGTGTGTGT GCCTGTGTGC TACGGAGAGT CCCGTATTCT CAGAGTAAAA
GTTGTTCTGG   480

AATGATCTCG CCAAAAAGGA CATCTTTGGA GCCAGTGATC CGTATGTGAA
ACTTTCATTG   540

TACGTAGCGG ATGAGAATAG AGAACTTGCT TTGGTCCAGA CAAAAACAAT
TAAAAAGACA   600

CTGAACCCAA AATGGAATGA AGAATTTTAT TTCAGGGTAA ACCCATCTAA
TCACAGACTC   660

CTATTTGAAG TATTTGACGA AAATAGACTG ACACGAGACG ACTTCCTGGG
CCAGGTGGAC   720

GTGCCCCTTA GTCACCTTCC GACAGAAGAT CCAACCATGG AGCGACCCTA
TACATTTAAG   780

GACTTTCTCC TCAGACCAAG AAGTCATAAG TCTCGAGTTA AGGGATTTTT
GCGATTGAAA   840

ATGGCCTATA TGCCAAAAAA TGGAGGTCAA GATGAAGAAA ACAGTGACCA
GAGGGATGAC   900

ATGGAGCATG GATGGGAAGT TGTTGACTCA AATGACTCGG CTTCTCAGCA
CCAAGAGGAA   960

CTTCCTCCTC CTCCTCTGCC TCCCGGGTGG GAAGAAAAAG TGGACAATTT
AGGCCGAACT   1020

TACTATGTCA ACCACAACAA CCGGACCACT CAGTGGCACA GACCAAGCCT
GATGGACGTG   1080

TCCTCGGAGT CGGACAATAA CATCAGACAG ATCAACCAGG AGGCAGCACA
CCGGCGCTTC   1140

CGCTCCCGCA GGCACATCAG CGAAGACTTG GAGCCCGAGC CCTCGGAGGG
CGGGGATGTC   1200

CCCGAGCCTT GGGAGACCAT TTCAGAGGAA GTGAATATCG CTGGAGACTC
TCTCGGTGTG   1260

GTTTTGCCCC CACCACCGGC CTCCCCAGGA TCTCGGACCA GCCCTCAGGA
GCTGTCAGAG   1320

GAACTAAGCA GAAGGCTTCA GATCACTCCA GACTCCAATG GGGAACAGTT
CAGCTCTTTG   1380

ATTCAAAGAG AACCCTCCTC AAGGTTGAGG TCATGCAGTG TCACCGACGC
AGTTGCAGAA   1440

CAGGGCCATC TACCACCGCC ATCAGTGGCC TATGTACATA CCACGCCGGG
TCTGCCTTCA   1500

GGCTGGGAAG AAAGAAAAGA TGCTAAGGGG CGCACATACT ATGTCAATCA
TAACAATCGA   1560

ACCACAACTT GGACTCGACC TATCATGCAG CTTGCAGAAG ATGGTGCGTC
CGGATCAGCC   1620

ACAAACAGTA ACAACCATCT AATCGAGCCT CAGATCCGCC GGCCTCGTAG
CCTCAGCTCG   1680

CCAACAGTAA CTTTATTGCC CCGCTGGAGG GTGCCAAGGA CTCACCCGTA
CGTCGGGCTG   1740

TGAAAGACAC CCTTTCCAAC CCACAGTCCC CACAGCCATC ACCTTACAAC
TCCCCCAAAC   1800

CACAACACAA AGTCACACAG AGCTTCTTGC CACCCGGCTG GGAAATGAGG
ATAGCGCCAA   1860

ACGGCCGGCC CTTCTTCATT GATCATAACA CAAAGACTAC AACCTGGGAA
GATCCACGTT   1920

TGAAATTTCC AGTACATATG CGGTCAAAGA CATCTTTAAA CCCCAATGAC
CTTGGCCCCC   1980

TTCCTCCTGG CTGGGAAGAA AGAATTCACT TGGATGGCCG AACGTTTTAT
ATTGATCATA   2040

ATAGCAAAAT TACTCAGTGG GAAGACCCAA GACTGCAGAA CCCAGCTATT
ACTGGTCCGG   2100

CTGTCCCTTA CTCCAGAGAA TTTAAGCAGA AATATGACTA CTTCAGGAAG
AAATTAAAGA   2160

AACCTGCTGA TATCCCCAAT AGGTTTGAAA TGAAACTTCA CAGAAATAAC
ATATTTGAAG   2220

AGTCCTATCG GAGAATTATG TCCGTGAAAA GACCAGATGT CCTAAAAGCT
AGACTGTGGA   2280

TTGAGTTTGA ATCAGAGAAA GGTCTTGACT ATGGGGGTGT GGCCAGAGAA
TGGTTCTTCT   2340

TACTGTCCAA AGAGATGTTC AACCCCTACT ACGGCCTCTT TGAGTACTCT
GCCACGGACA   2400

ACTACACCCT TCAGATCAAC CCTAATTCAG GCCTCTGTAA TGAGGATCAT
TTGTCCTACT   2460

TCACTTTTAT TGGAAGAGTT GCTGGTCTGG CCGTATTTCA TGGGAAGCTC
TTAGATGGTT   2520

TCTTCATTAG ACCATTTTAC AAGATGATGT TGGGAAAGCA GATAACCCTG
AATGACATGG   2580

AATCTGTGGA TAGTGAATAT TACAACTCTT TGAAATGGAT CCTGGAGAAT
GACCCTACTG   2640

AGCTGGACCT CATGTTCTGC ATAGACGAAG AAAACTTTGG ACAGACATAT
CAAGTGGATT   2700

TGAAGCCCAA TGGGTCAGAA ATAATGGTCA CAAATGAAAA CAAAAGGGAA
TATATCGACT   2760

TAGTCATCCA GTGGAGATTT GTGAACAGGG TCCAGAAGCA GATGAACGCC
TTCTTGGAGG   2820

GATTCACAGA ACTACTTCCT ATTGATTTGA TTAAAATTTT TGATGAAAAT
GAGCTGGAGT   2880

TGCTCATGTG CGGCCTCGGT GATGTGGATG TGAATGACTG GAGACAGCAT
TCTATTTACA   2940

AGAACGGCTA CTGCCCAAAC CACCCCGTCA TTCAGTGGTT CTGGAAGGCT
GTGCTACTCA   3000

TGGACGCCGA AAAGCGTATC CGGTTACTGC AGTTTGTCAC AGGGACATCG
CGAGTACCTA   3060

TGAATGGATT TGCCGAACTT TATGGTTCCA ATGGTCCTCA GCTGTTTACA
ATAGAGCAAT   3120

GGGGCAGTCC TGAGAAACTG CCCAGAGCTC ACACATGCTT TAATCGCCTT
GACTTACCTC   3180

CATATGAAAC CTTTGAAGAT TTACGAGAGA AACTTCTCAT GGCCGTGGAA
AATGCTCAAG   3240

GATTTGAAGG GGTGGATTAA GCACCCTGTG CCTCGGGGGT GGTTGTTCTT
CAAGCAAGTT   3300

CTGCTTGCAC TTTTGCATTT GCCTAACAGA CTTTTGCAGA GGCGATGGCA
GAGAGCAGCT   3360

GCAGGCATGG TCCCTGGAGC CGAGCCTTCA CCACGCACTC GTCCAAGTTC
GGGATGCGGG   3420

AACCTGGTCC CAGCTTGAGT TCCTGCCTTT CCCACCACAA ATTATCAACT
GGTTGATGTG   3480

TACACTAATT ACATTTCAGG AGGACTTAAT GCTATTTATG TTGTCCTCTG
CAGGCAAAGC   3540

CCTTAATAAA TATTTTACAT CCTTTCTAAT GACAATGAAT GGAATTAATC
ACTCAACAGG   3600

TATAGTATTA CGACTCATGT TTACTTTTTA AAATGATTTA GACCGATTTT
CAGATTTTAT   3660

TTCGTTATGA TTAAAGATGT CTCATGTACT TGGAAAAGTG AGCATTTTTT
TTTTTTTTTG   3720

TATTTCACTT TCATACCAGG CTTAATGTCA ATGACATTTT TATTTTTGAA
GTACTCTGAC   3780

ACCTCCACCC TCTACTTTAT TAGAATTGGA AGGCAAATTT TTGTCCAAAA
ACCTACAGAC   3840

AAGTACTTTG AGAGAATTTC CAATATAATA TTAGACATAA TGATAATTTT
TTCCATACTC   3900

AGAATGAAAA ACTGGATATT ACGTTTTTGT TTTGGGGTTT TTTTGTACAA
ATTTAGCTAA   3960

TAGCTACAGG CTGAGAGAAT TGTAACATAG CATGACAAAT TTTGTGTTGA
CTTGAAAGGA   4020

ATCACACCAT TATTCCTTAG AAGTAATTAC ATGTGTTCTA ACACATTTGA
GACAGGGTTG   4080

GACTCCCATT TCTCATCCGA GAAATTACTT AACCCTTCCT GGGCGCTGTA
CAGTCATCTT   4140

TTATTCTATT TCCTCTTTGC TGTTTGTAGT AGAGACATTT TGAATGAAAC
TTGGCACTGC   4200

TTGATTCAAA ACTGTGGAAA CCAGATCTGT TTAGTCTCCT GTTTGTATGC
GTTTGCTAAT   4260

GGTAGCTAAA TAACCAGTTT TTGTTGTAAA TGCACCAATT CTGAAGGCAC
TTTATGTACT   4320

ACATGGAGGT CATATCTGGT TTTGTTTTTA TTTTTTTATC ATGAACATTA
AATGTGATGA   4380

TGATTTCTTT TCCCTGCACA CATCTTTCCG GTGCAATATC TATCAATTGT
GAATCTGGCT   4440

GCTGGTGTAT AAAAACCTGG ATGTAAAGCT GAGCCTACAG ACCTGTCCTC
ACCAACTGTT   4500

TTGTGATTTC TACTCAACTA CAAAGATTTA TTTAATGTAC TCTTAATCTA
ACTGAGTTTT   4560

GTTACCAATG ACCTGTTGCA TGCTTCAATA CCGTGTACTG CCTGAGTTGT
GCCTCTTGTG   4620

TGCTAGATTA AAAGTGAGAC AGAGACTTGA CTTGATCCTC TGAGCCTCAA
GCTATTGAGC   4680

TGGTAGTGGC AGAGGACTGA GGGTACCTGC ACAGTTTGAT TCTTTTCCCA
CGTTGTAAGT   4740

CTCCATTGCA GAATTGTCGT GCGTTTGAGA AAACACCTGA GGCAGTGTGG
GAGTTGAACG   4800

ACCCTGCTGT CCTTTTTAAC CTGTGTTGTC CTAGACCTGT CGGGGCAGTC
AGGGGACACT   4860

AGAGATTTGA TCTCATGCGA GTCATCAATA GGACAAAAAA GTTGTGGTTT
GGGGAGGTCT   4920

GTTTGTTACA TAAAAAGGAC CTTTCGGTGT AAGAAATTGC CGTTTTTACC
CTGCCCTGGC   4980

TGGCATGTGA GAAGCCATGG AAGGTTGTGG TTGTAAATGA GTTGTCTAAA
GGGGTGCAGA   5040

GGCCTGAGGT TTCTAAAAGA AGGTAGATTT CTACAGAGCT GAGTGTTGGT
TCCTTTTTCT   5100

TATTGGTTGA AAATTACCTG GTAGTGATCA GAAAACTTAG ATGCTATGTA ACTC
5154

(2) INFORMATION FOR SEQ ID NO: 2:

   (i) SEQUENCE CHARACTERISTICS:
     (A) LENGTH: 975 amino acids
     (B) TYPE: amino acid
     (C) STRANDEDNESS: single
     (D) TOPOLOGY: linear

   (ii) MOLECULE TYPE: peptide

   (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

Met Phe Arg Leu Arg Ser Trp Ala Ser Ser Thr Thr Gly Ser Arg Tyr
1   5   10   15

Gly Ser Ala Phe Cys Gly Ser Pro Thr Leu Ala Trp Cys Val Cys Val
     20   25   30

Pro Val Cys Tyr Gly Glu Ser Arg Ile Leu Arg Val Lys Val Val Ser
     35   40   45

Gly Ile Asp Leu Ala Lys Lys Asp Ile Phe Gly Ala Ser Asp Pro Tyr
   50   55   60

Val Lys Leu Ser Leu Tyr Val Ala Asp Glu Asn Arg Glu Leu Ala Leu
65   70   75   80

Val Gln Thr Lys Thr Ile Lys Lys Thr Leu Asn Pro Lys Trp Asn Glu
     85   90   95

Glu Phe Tyr Phe Arg Val Asn Pro Ser Asn His Arg Leu Leu Phe Glu
    100   105   110

Val Phe Asp Glu Asn Arg Leu Thr Arg Asp Asp Phe Leu Gly Gln Val
   115   120   125

Asp Val Pro Leu Ser His Leu Pro Thr Glu Asp Pro Thr Met Glu Arg
  130   135   140

Pro Tyr Thr Phe Lys Asp Phe Leu Leu Arg Pro Arg Ser His Lys Ser
145   150   155   160

Arg Val Lys Gly Phe Leu Arg Leu Lys Met Ala Tyr Met Pro Lys Asn
    165   170   175

Gly Gly Gln Asp Glu Glu Asn Ser Asp Gln Arg Asp Asp Met Glu His
    180   185   190

Gly Trp Glu Val Val Asp Ser Asn Asp Ser Ala Ser Gln His Gln Glu
    195   200   205

Glu Leu Pro Pro Pro Pro Leu Pro Pro Gly Trp Glu Glu Lys Val Asp
  210   215   220

Asn Leu Gly Arg Thr Tyr Tyr Val Asn His Asn Asn Arg Thr Thr Gln
225   230   235   240

Trp His Arg Pro Ser Leu Met Asp Val Ser Ser Glu Ser Asp Asn Asn
    245   250   255

Ile Arg Gln Ile Asn Gln Glu Ala Ala His Arg Arg Phe Arg Ser Arg
    260   265   270

Arg His Ile Ser Glu Asp Leu Glu Pro Glu Pro Ser Glu Gly Gly Asp
    275   280   285

Val Pro Glu Pro Trp Glu Thr Ile Ser Glu Glu Val Asn Ile Ala Gly
  290   295   300

Asp Ser Leu Gly Val Val Leu Pro Pro Pro Ala Ser Pro Gly Ser
305   310   315   320

Arg Thr Ser Pro Gln Glu Leu Ser Glu Glu Leu Ser Arg Arg Leu Gln
    325   330   335

Ile Thr Pro Asp Ser Asn Gly Glu Gln Phe Ser Ser Leu Ile Gln Arg
    340   345   350

Glu Pro Ser Ser Arg Leu Arg Ser Cys Ser Val Thr Asp Ala Val Ala
   355   360   365

Glu Gln Gly His Leu Pro Pro Pro Ser Val Ala Tyr Val His Thr Thr
   370   375   380

Pro Gly Leu Pro Ser Gly Trp Glu Glu Arg Lys Asp Ala Lys Gly Arg
385   390   395   400

Thr Tyr Tyr Val Asn His Asn Asn Arg Thr Thr Thr Trp Thr Arg Pro
            405   410   415

Ile Met Gln Leu Ala Glu Asp Gly Ala Ser Gly Ser Ala Thr Asn Ser
            420   425   430

Asn Asn His Leu Ile Glu Pro Gln Ile Arg Arg Pro Arg Ser Leu Ser
   435   440   445

Ser Pro Thr Val Thr Leu Xaa Ala Pro Leu Glu Gly Ala Lys Asp Ser
   450   455   460

Pro Val Arg Arg Ala Val Lys Asp Thr Leu Ser Asn Pro Gln Ser Pro
465   470   475   480

Gln Pro Ser Pro Tyr Asn Ser Pro Lys Pro Gln His Lys Val Thr Gln
            485   490   495

Ser Phe Leu Pro Pro Gly Trp Glu Met Arg Ile Ala Pro Asn Gly Arg
            500   505   510

Pro Phe Phe Ile Asp His Asn Thr Lys Thr Thr Thr Trp Glu Asp Pro
            515   520   525

Arg Leu Lys Phe Pro Val His Met Arg Ser Lys Thr Ser Leu Asn Pro
   530   535   540

Asn Asp Leu Gly Pro Leu Pro Pro Gly Trp Glu Glu Arg Ile His Leu
545   550   555   560

Asp Gly Arg Thr Phe Tyr Ile Asp His Asn Ser Lys Ile Thr Gln Trp
            565   570   575

Glu Asp Pro Arg Leu Gln Asn Pro Ala Ile Thr Gly Pro Ala Val Pro
            580   585   590

Tyr Ser Arg Glu Phe Lys Gln Lys Tyr Asp Tyr Phe Arg Lys Lys Leu
   595   600   605

Lys Lys Pro Ala Asp Ile Pro Asn Arg Phe Glu Met Lys Leu His Arg
    610   615   620

Asn Asn Ile Phe Glu Glu Ser Tyr Arg Arg Ile Met Ser Val Lys Arg
625   630   635   640

Pro Asp Val Leu Lys Ala Arg Leu Trp Ile Glu Phe Glu Ser Glu Lys
         645   650   655

Gly Leu Asp Tyr Gly Gly Val Ala Arg Glu Trp Phe Phe Leu Leu Ser
         660   665   670

Lys Glu Met Phe Asn Pro Tyr Tyr Gly Leu Phe Glu Tyr Ser Ala Thr
    675   680   685

Asp Asn Tyr Thr Leu Gln Ile Asn Pro Asn Ser Gly Leu Cys Asn Glu
    690   695   700

Asp His Leu Ser Tyr Phe Thr Phe Ile Gly Arg Val Ala Gly Leu Ala
705   710   715   720

Val Phe His Gly Lys Leu Leu Asp Gly Phe Phe Ile Arg Pro Phe Tyr
         725   730   735

Lys Met Met Leu Gly Lys Gln Ile Thr Leu Asn Asp Met Glu Ser Val
         740   745   750

Asp Ser Glu Tyr Tyr Asn Ser Leu Lys Trp Ile Leu Glu Asn Asp Pro
         755   760   765

Thr Glu Leu Asp Leu Met Phe Cys Ile Asp Glu Glu Asn Phe Gly Gln
         770   775   780

Thr Tyr Gln Val Asp Leu Lys Pro Asn Gly Ser Glu Ile Met Val Thr
785   790   795   800

Asn Glu Asn Lys Arg Glu Tyr Ile Asp Leu Val Ile Gln Trp Arg Phe
         805   810   815

Val Asn Arg Val Gln Lys Gln Met Asn Ala Phe Leu Glu Gly Phe Thr
         820   825   830

Glu Leu Leu Pro Ile Asp Leu Ile Lys Ile Phe Asp Glu Asn Glu Leu
         835   840   845

Glu Leu Leu Met Cys Gly Leu Gly Asp Val Asp Val Asn Asp Trp Arg
    850   855   860

Gln His Ser Ile Tyr Lys Asn Gly Tyr Cys Pro Asn His Pro Val Ile
865   870   875   880

Gln Trp Phe Trp Lys Ala Val Leu Leu Met Asp Ala Glu Lys Arg Ile
             885   890   895

Arg Leu Leu Gln Phe Val Thr Gly Thr Ser Arg Val Pro Met Asn Gly
             900   905   910

Phe Ala Glu Leu Tyr Gly Ser Asn Gly Pro Gln Leu Phe Thr Ile Glu
             915   920   925

Gln Trp Gly Ser Pro Glu Lys Leu Pro Arg Ala His Thr Cys Phe Asn
             930   935   940

Arg Leu Asp Leu Pro Pro Tyr Glu Thr Phe Glu Asp Leu Arg Glu Lys
945   950   955   960

Leu Leu Met Ala Val Glu Asn Ala Gln Gly Phe Glu Gly Val Asp
             965   970   975

(2) INFORMATION FOR SEQ ID NO: 3:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 854 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: other nucleic acid

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

ACAATGGGGG CGTGGCAGAG AATGGTTCTT CTTACTGTCC AAAGAGATGT
TTAACCCCTA   60

CTATGGCCTC TTCGAGTACT CTGCCACGGA CAACTACACA CTTCAGATCA
ATCCCAACTC   120

AGGCCTCTGT AATGAAGACC ATTTGTCCTA TTTCACCTTC ATTGGAAGAG
TTGCTGGCCT   180

AGCGGTGTTT CATGGGAAAC TCTTAGATGG ATTCTTCATT CGACCATTCT
ACAAGATGAT   240

GCTGGGGAAG CAGATAACGC TGAACGACAT GGAGTCCGTG GACAGCGAGT
ACTACAACTC   300

TTTGAAGTGG ATCTTAGAAA ACGACCCCAC GGAACTTGAC CTCATGTTCT
GCATAGACGA   360

GAGAACTTTG GGCAGACATA CCAAGTGGAT CTGAAGCCCA ACGGGTCAGA
AATAATGGTA   420

ACCAATGAGA ACAAACGAGA ATACATTGAC TTAGTCATCC AGTGGAGATT
TGTGAACAGG   480

GTCCAGAAGC AAATGAATGC CTTCTTGGAG GGATTTACAG AACTTCTTCC
AATCGACTTG   540

ATTAAAATTT TTGATGAAAA TGAGCTGGAG TTGCTGATGT GCGGCCTTGG
TGATGTCGAC   600

GTGAACGACT GGAGACAGCA CTCTATTTAC AAGAACGGCT ACTGCCCCAA
CCACCCTGTC   660

ATCCAGTGGT TCTGGAAGGC CGTGCTCCTG ATGGATGCTG AGAAGCGCAT
CCGGTTACTA   720

CAGTTTGTCA CAGGCACCTC CAGAGTACCC ATGAATGGAT TTGCCGAACT
CTATGGTTCC   780

AATGGTCCTC AGCTGTTTAC AATAGAGCAA TGGGGCAGTC CGAAAAACTA
CCAGAGCTCT   840

ACATGCTTAA TCGC   854

(2) INFORMATION FOR SEQ ID NO: 4:

   (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 604 amino acids
       (B) TYPE: amino acid
       (C) STRANDEDNESS: single
       (D) TOPOLOGY: linear

   (ii) MOLECULE TYPE: peptide

   (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

His Ala Cys Ser Asn Ala Ala Ser Arg Ala Ala Ala Arg Val Ala Ala
1   5   10   15

Arg Cys Thr Ala Arg Ser Arg Ser Gly Arg Arg Ser Ser Ser Val Ser
          20   25   30

Arg Ser Ser Ser Arg Gly Ala Ser Ser Ser Met Ser Ser Asp Met Ala
          35   40   45

Ala Asp Ser Ala Val Ser Asp Val Trp Cys Asp Lys Thr Asp Gly Gly
          50   55   60

Gly Ser Gly Ser Asp Val Thr Asp Thr Cys Cys Gly Cys Trp Asn Asn
65   70   75   80

Ser His Val Thr Ala Asp Tyr His Asn Asp Asp Thr Arg Val Val Arg
               85   90   95

Val Lys Val Ala Gly Gly Ala Lys Lys Asp Gly Ala Ser Asp Tyr Val
               100   105   110

Arg Val Thr Tyr Asp Met Ser Gly Thr Ser Val Thr Lys Thr Lys Lys
          115   120   125

Ser Asn Lys Trp Asn Arg Val Arg His Arg Val Asp Asn Arg Thr Arg
          130   135   140

Asp Asp Gly Val Asp Val Tyr Thr Asn Arg Met Arg Tyr Thr Lys Asp
145   150   155   160

Val His Arg Ser His Lys Ser Arg Val Lys Gly Tyr Arg Lys Met Thr
               165   170   175

Tyr Lys Asn Gly Ser Asp Asn Ala Asp Ala Gly Trp Val Val Asp Asp
               180   185   190

Ala Ala Thr His His Ser Gly Trp Arg Asp Val Gly Arg Thr Tyr Tyr
          195   200   205

Val Asn His Ser Arg Arg Thr Trp Lys Arg Ser Asp Asp Asp Thr Asp
          210   215   220

Asp Asn Asp Asp Met Ala Arg Ala Thr Thr Arg Arg Ser Asp Val Asp
225   230   235   240

Gly Asp Asn Arg Ser Asn Trp Val Arg Asp Asn Thr Tyr Ser Gly Ala
               245   250   255

Val Ser Ser Gly His Asp Val Thr His Ala Asn Thr Arg Ala Val Cys
     260   265   270

Gly Asn Ala Thr Ser Val Thr Ser Ser Asn His Ser Ser Arg Gly Gly
     275   280   285

Ser Thr Cys Thr Val Thr Ser Ser Gly Gly Trp Lys Asp Asp Arg Gly
     290   295   300

Arg Ser Tyr Tyr Val Asp His Asn Ser Lys Thr Thr Thr Trp Ser Lys
305   310   315   320

Thr Met Asp Asp Arg Ser Lys Ala His Arg Gly Lys Thr Asp Ser Asn
        325   330   335

Asp Gly Gly Trp Arg Thr His Thr Asp Gly Arg Val Asn His Asn Lys
     340   345   350

Lys Thr Trp Asp Arg Asn Val Ala Thr Gly Ala Val Tyr Ser Arg Asp
     355   360   365

Tyr Lys Arg Lys Tyr Arg Arg Lys Lys Lys Thr Asp Asn Lys Met Lys
     370   375   380

Arg Arg Ala Asn Asp Ser Tyr Arg Arg Met Gly Val Lys Arg Ala Asp
385   390   395   400

Lys Ala Arg Trp Asp Gly Lys Gly Asp Tyr Gly Gly Val Ala Arg Trp
     405   410   415

Ser Lys Met Asn Tyr Tyr Gly Tyr Ser Ala Thr Asp Asn Tyr Thr Asn
     420   425   430

Asn Ser Gly Cys Asn Asp His Ser Tyr Lys Gly Arg Val Ala Gly Met
     435   440   445

Ala Val Tyr His Gly Lys Asp Gly Arg Tyr Lys Met Met Lys Thr His
     450   455   460

Asp Met Ser Val Asp Ser Tyr Tyr Ser Ser Arg Trp Asn Asp Thr Asp
465   470   475   480

Arg Asp Gly Thr His His Lys Thr Gly Gly Ser Val Val Thr Asn Lys
        485   490   495

Asn Lys Lys Tyr Tyr Val Trp Arg Val Asn Arg Lys Met Ala Ala Lys
     500   505   510

Gly Asp Lys Asp Asn Met Cys Gly Gly Asp Val Asp Val Asn Asp Trp
    515   520   525

Arg His Thr Lys Tyr Lys Asn Gly Tyr Ser Met Asn His Val His Trp
  530   535   540

Trp Lys Ala Val Trp Met Met Asp Ser Lys Arg Arg Val Thr Gly Thr
545   550   555   560

Xaa Ser Arg Val Met Asn Gly Ala Tyr Gly Ser Asn Gly Ser Thr Val
    565   570   575

Trp Gly Thr Asp Lys Arg Ala His Thr Cys Asn Arg Asp Tyr Ser Asp
   580   585   590

Trp Asp Lys Met Ala Asn Thr Gly Asp Gly Val Asp
  595   600

(2) INFORMATION FOR SEQ ID NO: 5:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 615 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: other nucleic acid

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

TGCTGCAAGT GACAGGTTCC AAGAAGCCCG AGGGCTCAGA GCTGAATGAT
GAAGCGCAGT   60

CCCCAAAGTG CCTGGCCACC CCTCCCTCCC TGGATCACTG CTGCCTGGGC
TTGATTGATT   120

GATTGATTGA TTGATTGATT GATTTGAGA GAGATTCTCA CTGTCACCCA
GGCTGGAGTA   180

CAGTGGTGCG ATCTCGGCTC ACTGCAGCCT CTGCCTCCCG GGTTCAAGCA
ATTCTCCTGC   240

CTCAGCCTCC CAAGTAGCTG GGACTACAGG CACGCGCCAC CACACCCAGC
TAATTTTGTA   300

-16-

TTTTTAGTAA AAGACGGGGT TTCACCATGT TGGGCCAGGA TGGTCTTGAT
CTCCTGACCT   360

CATGATCCAC CCGCCCCGGC TTCCAAAGTG CTGGGATACA GGCATGAACC
CGACGCGCCC   420

AGCATGGACA TTTTTTTTTA ATCCCCTGCC CTTTTCTTGG GCATAATTCA
TTGCAGGTCT   480

CTTCTATACA GATCATGGAA AACACATTTT CTTAACTGAG TTTTATTATT
TATACCCAGC   540

ACCTCATGAC ATTTACCCTG TTACAACAAA ATGGGCACCT GCCAAAACAA
CTTTATATAA   600

GGATGCTCCA GGCCT   615